



wwPDB X-ray Structure Validation Summary Report

Sep 5, 2017 – 06:55 AM EDT

PDB ID : 4UF6
Title : UCH-L5 in complex with ubiquitin-propargyl bound to an activating fragment of INO80G
Authors : Sahtoe, D.D.; Van Dijk, W.J.; El Oualid, F.; Ekkebus, R.; Ovaa, H.; Sixma, T.K.
Deposited on : unknown
Resolution : 3.69 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the  symbol.

The following versions of software and data (see [references](#) ) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.9-1692
EDS : rb-20029824
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20029824

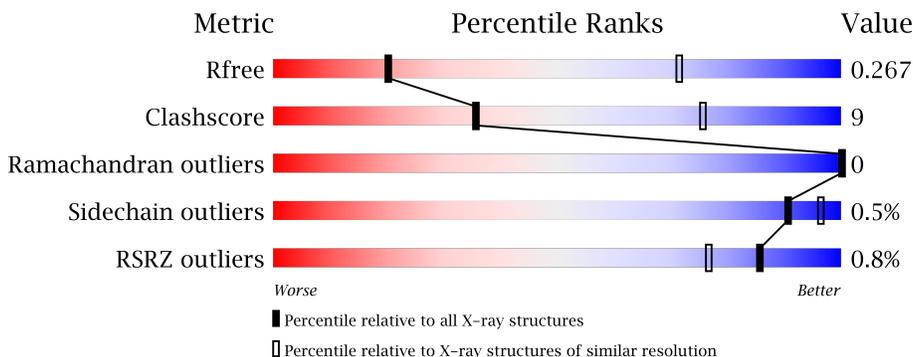
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.69 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	1234 (3.90-3.50)
Clashscore	112137	1377 (3.90-3.50)
Ramachandran outliers	110173	1323 (3.90-3.50)
Sidechain outliers	110143	1320 (3.90-3.50)
RSRZ outliers	101464	1262 (3.90-3.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	331	 4% 66% 22% 12%
1	D	331	 69% 20% 11%
1	G	331	 4% 66% 21% 13%
1	J	331	 68% 21% 11%
2	B	76	 4% 80% 18% 1%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	76	 87% 13%
2	H	76	 91% 9%
2	K	76	 83% 17%
3	C	65	 2% 52% 25% 23%
3	F	65	 3% 58% 18% 23%
3	I	65	 2% 55% 20% 25%
3	L	65	 2% 52% 23% 25%

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 13486 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	292	Total 2360	C 1503	N 405	O 440	S 12	0	0	0
1	D	295	Total 2382	C 1515	N 409	O 446	S 12	0	0	0
1	G	287	Total 2313	C 1475	N 397	O 430	S 11	0	0	0
1	J	294	Total 2377	C 1514	N 407	O 444	S 12	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP Q9Y5K5
A	-1	PRO	-	expression tag	UNP Q9Y5K5
A	0	GLY	-	expression tag	UNP Q9Y5K5
D	-2	GLY	-	expression tag	UNP Q9Y5K5
D	-1	PRO	-	expression tag	UNP Q9Y5K5
D	0	GLY	-	expression tag	UNP Q9Y5K5
G	-2	GLY	-	expression tag	UNP Q9Y5K5
G	-1	PRO	-	expression tag	UNP Q9Y5K5
G	0	GLY	-	expression tag	UNP Q9Y5K5
J	-2	GLY	-	expression tag	UNP Q9Y5K5
J	-1	PRO	-	expression tag	UNP Q9Y5K5
J	0	GLY	-	expression tag	UNP Q9Y5K5

- Molecule 2 is a protein called POLYUBIQUITIN-B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	76	Total 601	C 379	N 105	O 116	S 1	0	0	0
2	E	76	Total 601	C 379	N 105	O 116	S 1	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	76	Total	C	N	O	S	0	0	0
			601	379	105	116	1			
2	K	76	Total	C	N	O	S	0	0	0
			601	379	105	116	1			

- Molecule 3 is a protein called NUCLEAR FACTOR RELATED TO KAPPA-B-BINDING PROTEIN.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	50	Total	C	N	O	0	0	0
			417	267	64	86			
3	F	50	Total	C	N	O	0	0	0
			417	267	64	86			
3	I	49	Total	C	N	O	0	0	0
			408	262	63	83			
3	L	49	Total	C	N	O	0	0	0
			408	262	63	83			

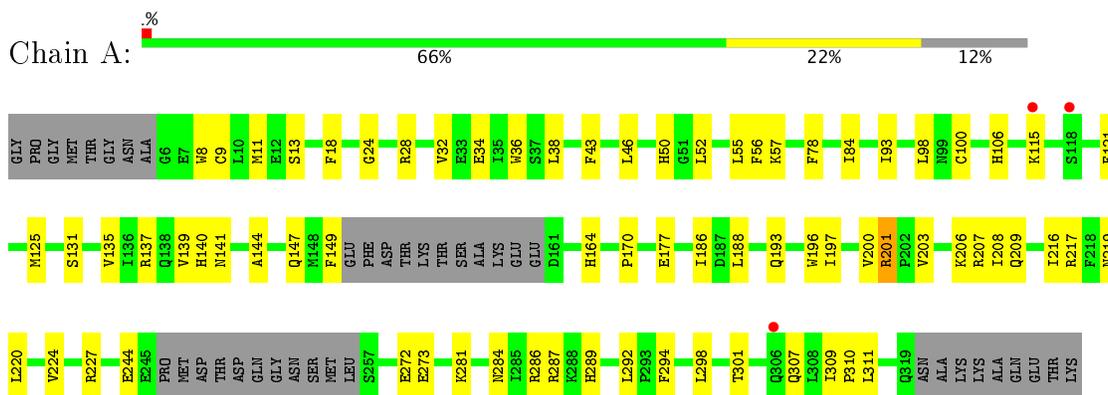
There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	37	GLY	-	expression tag	UNP Q6P4R8
C	38	PRO	-	expression tag	UNP Q6P4R8
C	39	GLY	-	expression tag	UNP Q6P4R8
F	37	GLY	-	expression tag	UNP Q6P4R8
F	38	PRO	-	expression tag	UNP Q6P4R8
F	39	GLY	-	expression tag	UNP Q6P4R8
I	37	GLY	-	expression tag	UNP Q6P4R8
I	38	PRO	-	expression tag	UNP Q6P4R8
I	39	GLY	-	expression tag	UNP Q6P4R8
L	37	GLY	-	expression tag	UNP Q6P4R8
L	38	PRO	-	expression tag	UNP Q6P4R8
L	39	GLY	-	expression tag	UNP Q6P4R8

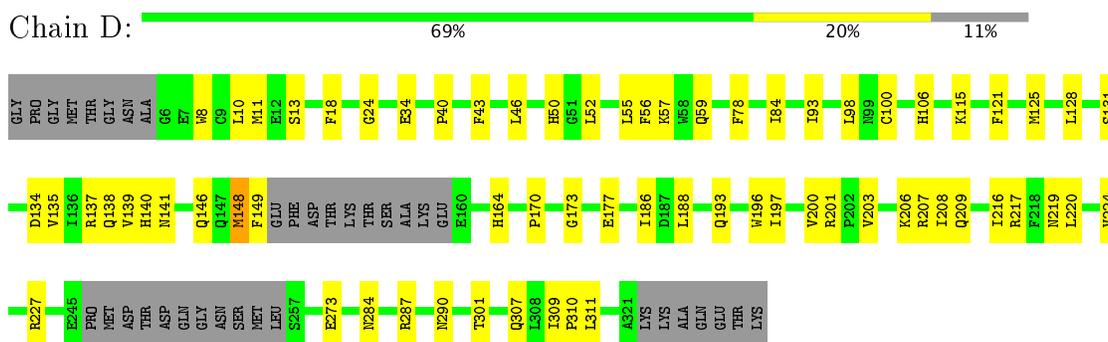
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ($RSRZ > 2$). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

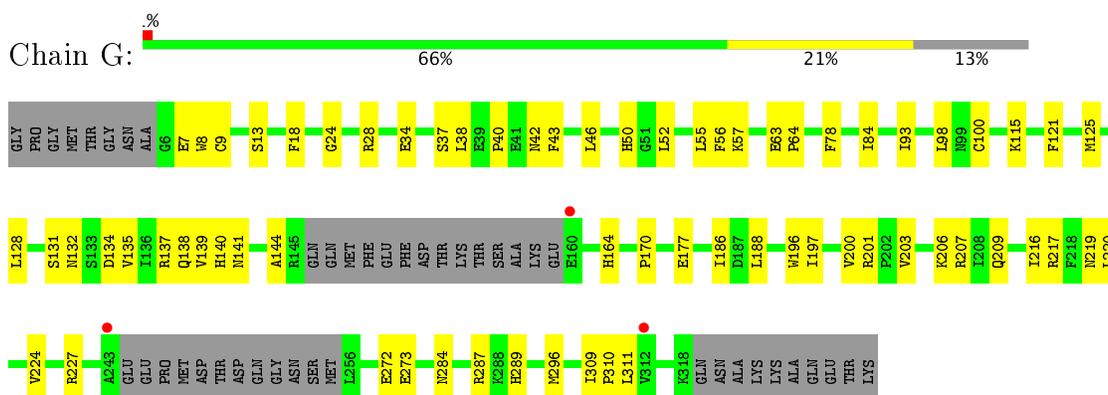
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L5



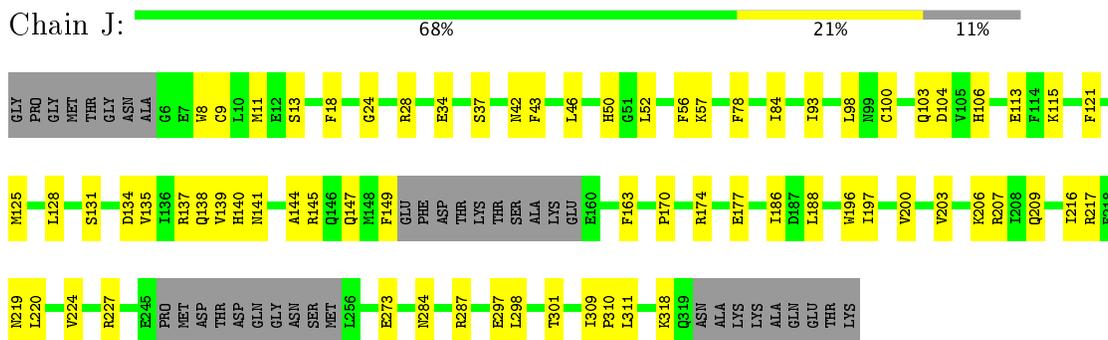
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L5



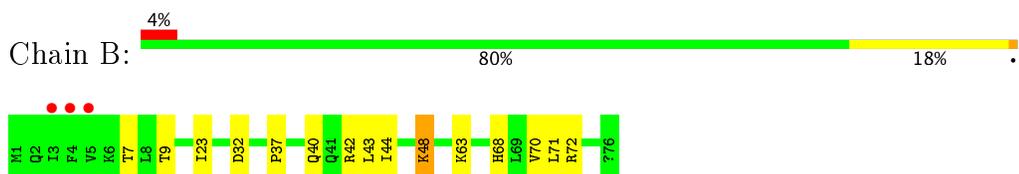
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L5



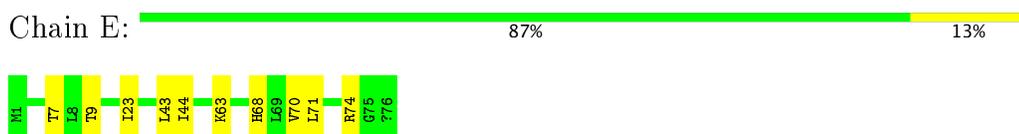
- Molecule 1: UBIQUITIN CARBOXYL-TERMINAL HYDROLASE ISOZYME L5



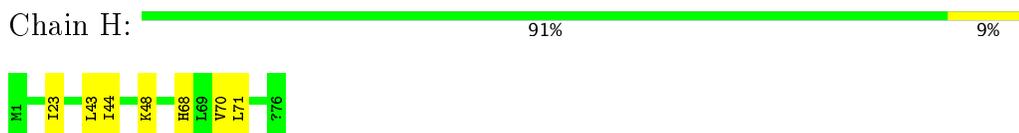
- Molecule 2: POLYUBIQUITIN-B



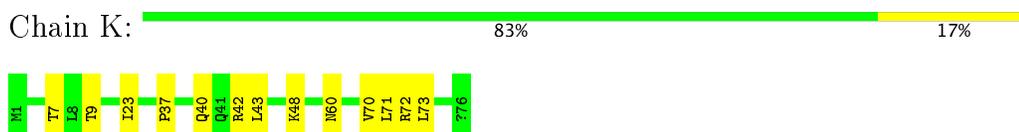
- Molecule 2: POLYUBIQUITIN-B



- Molecule 2: POLYUBIQUITIN-B



- Molecule 2: POLYUBIQUITIN-B



- Molecule 3: NUCLEAR FACTOR RELATED TO KAPPA-B-BINDING PROTEIN



- Molecule 3: NUCLEAR FACTOR RELATED TO KAPPA-B-BINDING PROTEIN





- Molecule 3: NUCLEAR FACTOR RELATED TO KAPPA-B-BINDING PROTEIN



- Molecule 3: NUCLEAR FACTOR RELATED TO KAPPA-B-BINDING PROTEIN



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	152.08Å 137.79Å 98.92Å 90.00° 102.54° 90.00°	Depositor
Resolution (Å)	100.99 – 3.69 47.74 – 3.69	Depositor EDS
% Data completeness (in resolution range)	97.9 (100.99-3.69) 98.0 (47.74-3.69)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.43 (at 3.67Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.238 , 0.269 0.240 , 0.267	Depositor DCC
R_{free} test set	1059 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	82.8	Xtriage
Anisotropy	0.568	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 29.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	13486	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.54% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: AYE

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >5	RMSZ	# Z >5
1	A	0.44	0/2404	0.64	2/3236 (0.1%)
1	D	0.44	0/2426	0.62	0/3266
1	G	0.41	0/2356	0.59	0/3173
1	J	0.46	0/2421	0.62	0/3259
2	B	0.34	0/603	0.57	0/811
2	E	0.39	0/603	0.61	0/811
2	H	0.35	0/603	0.57	0/811
2	K	0.35	0/603	0.57	0/811
3	C	0.43	0/427	0.63	0/581
3	F	0.45	0/427	0.68	0/581
3	I	0.43	0/418	0.63	0/569
3	L	0.46	0/418	0.63	0/569
All	All	0.43	0/13709	0.61	2/18478 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	201	ARG	NE-CZ-NH1	7.37	123.99	120.30
1	A	201	ARG	NE-CZ-NH2	-6.35	117.13	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen

atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2360	0	2348	62	1
1	D	2382	0	2365	54	0
1	G	2313	0	2311	53	0
1	J	2377	0	2365	50	1
2	B	601	0	630	12	2
2	E	601	0	630	10	1
2	H	601	0	630	6	0
2	K	601	0	630	8	1
3	C	417	0	383	17	0
3	F	417	0	383	11	0
3	I	408	0	377	7	1
3	L	408	0	377	11	1
All	All	13486	0	13429	247	4

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 9.

The worst 5 of 247 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:40:PRO:HB3	1:G:201:ARG:HH22	1.20	1.02
1:D:135:VAL:O	1:D:139:VAL:HG23	1.63	0.99
1:A:135:VAL:O	1:A:139:VAL:HG23	1.62	0.98
1:J:135:VAL:O	1:J:139:VAL:HG23	1.62	0.97
1:G:135:VAL:O	1:G:139:VAL:HG23	1.62	0.97

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:244:GLU:O	1:J:174:ARG:NH1[2_755]	1.99	0.21
2:B:32:ASP:O	2:K:60:ASN:ND2[4_655]	2.05	0.15
2:E:63:LYS:NZ	3:I:82:ALA:O[4_656]	2.14	0.06
2:B:63:LYS:NZ	3:L:79:GLU:OE1[4_655]	2.18	0.02

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	286/331 (86%)	273 (96%)	13 (4%)	0	100	100
1	D	289/331 (87%)	276 (96%)	13 (4%)	0	100	100
1	G	281/331 (85%)	269 (96%)	12 (4%)	0	100	100
1	J	288/331 (87%)	276 (96%)	12 (4%)	0	100	100
2	B	73/76 (96%)	71 (97%)	2 (3%)	0	100	100
2	E	73/76 (96%)	70 (96%)	3 (4%)	0	100	100
2	H	73/76 (96%)	71 (97%)	2 (3%)	0	100	100
2	K	73/76 (96%)	71 (97%)	2 (3%)	0	100	100
3	C	48/65 (74%)	41 (85%)	7 (15%)	0	100	100
3	F	48/65 (74%)	41 (85%)	7 (15%)	0	100	100
3	I	47/65 (72%)	39 (83%)	8 (17%)	0	100	100
3	L	47/65 (72%)	39 (83%)	8 (17%)	0	100	100
All	All	1626/1888 (86%)	1537 (94%)	89 (6%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains [i](#)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	257/288 (89%)	257 (100%)	0	100	100
1	D	259/288 (90%)	257 (99%)	2 (1%)	85	93

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	252/288 (88%)	252 (100%)	0	100	100
1	J	259/288 (90%)	258 (100%)	1 (0%)	93	97
2	B	68/68 (100%)	67 (98%)	1 (2%)	70	88
2	E	68/68 (100%)	68 (100%)	0	100	100
2	H	68/68 (100%)	68 (100%)	0	100	100
2	K	68/68 (100%)	67 (98%)	1 (2%)	70	88
3	C	48/59 (81%)	48 (100%)	0	100	100
3	F	48/59 (81%)	48 (100%)	0	100	100
3	I	47/59 (80%)	46 (98%)	1 (2%)	59	84
3	L	47/59 (80%)	46 (98%)	1 (2%)	59	84
All	All	1489/1660 (90%)	1482 (100%)	7 (0%)	91	96

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	I	46	PRO
3	L	46	PRO
1	J	145	ARG
1	D	59	GLN
2	K	48	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	106	HIS
1	D	42	ASN
1	D	233	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled '#RSRZ > 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q < 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2	OWAB(Å ²)	Q < 0.9
1	A	292/331 (88%)	-0.09	3 (1%) 82 72	53, 79, 143, 181	0
1	D	295/331 (89%)	-0.13	0 100 100	48, 66, 125, 169	0
1	G	287/331 (86%)	-0.12	3 (1%) 82 72	50, 77, 146, 185	0
1	J	294/331 (88%)	-0.17	0 100 100	42, 63, 138, 170	0
2	B	75/76 (98%)	0.28	3 (4%) 39 28	107, 140, 187, 199	0
2	E	75/76 (98%)	-0.03	0 100 100	69, 86, 103, 106	0
2	H	75/76 (98%)	0.08	0 100 100	76, 90, 100, 102	0
2	K	75/76 (98%)	0.03	0 100 100	58, 76, 92, 97	0
3	C	50/65 (76%)	0.08	1 (2%) 65 53	102, 114, 126, 132	0
3	F	50/65 (76%)	0.21	2 (4%) 39 28	99, 113, 127, 133	0
3	I	49/65 (75%)	0.23	1 (2%) 65 53	106, 124, 144, 154	0
3	L	49/65 (75%)	0.33	1 (2%) 65 53	91, 115, 149, 158	0
All	All	1666/1888 (88%)	-0.05	14 (0%) 86 77	42, 80, 146, 199	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	71	GLN	3.9
1	G	243	ALA	3.8
3	F	81	SER	3.8
3	C	71	GLN	3.0
3	L	71	GLN	2.7

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.